



Metodología Encuesta Lechera 2019

INALE - MGAP

La Encuesta Lechera se realizó durante el año 2020 sobre una muestra representativa de 358 explotaciones con producción de leche con destino comercial. Se relevaron datos relativos al ejercicio 2018/2019 (1/07/2018 - 30/06/2019). Se realizó en forma telefónica.

Setiembre 2021



Equipo de trabajo

La Encuesta Lechera fue realizada en 2020 fue coordinada por el Instituto Nacional de la Leche (INALE) y la Oficina de Programación y Política Agropecuaria del Ministerio de Agricultura, Ganadería y Pesca (OPYPA- MGAP).

El formulario de la encuesta fue diseñado por el equipo técnico de INALE, representado por los ingenieros agrónomos Jorge Artagaveytia y Ana Pedemonte, con apoyo del OPYPA, integrado por los economistas Verónica Durán y Juan Baraldo, teniendo como base la encuesta lechera 2014 y consultando en ciertas áreas a referentes en las temáticas.

El diseño de la muestra estuvo a cargo del equipo de OPYPA contando con la base de las declaraciones de DICOSE 2018 como marco muestral.

El trabajo de campo fue realizado por Opción Consultores, se contó con encuestadores con conocimiento del sector (ingenieros agrónomos y veterinarios). Comenzó a realizarse de forma presencial pero dada la crisis sanitaria generada por la Covid-19 la mayor parte de las encuestas se realizaron de forma telefónica.

El equipo técnico de INALE realizó talleres de capacitación a los encuestadores sobre las preguntas en conjunto con Opción Consultores que se encargó de la operativa del llenado del formulario en tabletas o celulares.

El procesamiento de datos, publicación en la web y redacción del presente informe fue realizado por el equipo técnico de INALE que trabajó directamente en la encuesta con apoyo de OPYPA.



Definiciones previas

Se denomina **explotación con actividad lechera comercial** a toda superficie de tierra que se destina total o parcialmente a la producción lechera y que es manejada por una misma administración que por lo general cuenta con los mismos medios de producción tales como: mano de obra, maquinaria, instalaciones y animales de trabajo (con uno o más rodeos) y que puede estar concentrada en un lugar o localizada en diferentes puntos.

Se consideran como establecimientos remitentes a los que producen leche que se destina únicamente a la industria y queserías artesanales las que destinan parte de su leche a elaboración de productos lácteos en sus establecimientos (los que a su vez pueden también remitir parte de la leche a industrias).

El período relevado se corresponde con el ejercicio 2018/2019 comprendido entre el 1/7/2018 y el 30/6/2019.



Marco muestral

El marco muestral se construyó con los números de DICOSE de la declaración 2018, agrupando los DICOSE de una misma empresa por número de RUT, o por su identificador en el censo 2011 en los casos que no se dispone de número de RUT, en último lugar, en caso de no agrupar por los anteriores criterios, se consideró cada DICOSE como una empresa.

Se excluyeron del universo elegible los predios con menos de 10 vacas masa (156 individuos). Al mismo tiempo, se restringió el marco a aquellos establecimientos de los departamentos de Canelones, Colonia, Flores, Florida, Paysandú, San José, Soriano y Río Negro (y algunos casos de Maldonado).

El marco se estratificó por cantidad de vacas masa según los rangos que se muestran en el cuadro 1.

Cuadro 1. Estratificación de la población de interés

estratos VM		
Estrato	rango	cantidad de individuos
1	más de 1000 VM	23
2	301 a 1000 VM	252
3	126 a 300 VM	584
4	51 a 125 VM	1044
5	10 a 50 VM	1059

Además de dicha estratificación, se construyeron tres sub-poblaciones diferentes:

- Queseros
- Colonos
- Otros remitentes¹

De esta forma a priori quedan definidos 15 estratos, aunque no necesariamente todos los estratos tienen individuos.



Tamaño y asignación de la muestra

Todos los casos incluidos en el estrato 1 y los queseros del estrato 2 se incluyeron forzosamente, esto es, todos los individuos del estrato se asignaron a la muestra con probabilidad 1. En total resultaron 30 casos forzosos en la muestra original.

Por argumentos presupuestales originalmente se determinó que el tamaño de la muestra sería de 340 individuos a encuestar.

La determinación del tamaño de muestra en cada estrato se realizó con el criterio de Neyman, el cual pondera la cantidad de individuos de un estrato, la variabilidad estimada en el estrato de una variable relevante y conocida a priori (en este caso producción de leche por hectárea según la información de marco) y el costo de realizar una encuesta en cada estrato que se supuso era igual entre estratos.

$$n_{0h} = n_0 * \frac{(N_h * s_h / \sqrt{c})}{\sum_h (N_h * s_h / \sqrt{c})}$$

Donde n_{0h} es el tamaño muestral óptimo del estrato h, n_0 es el tamaño muestral deseado, N_h es el tamaño de la población perteneciente al estrato h, s_h es el desvío estándar de la variable relevante y c es el costo de una encuesta.

Teniendo en cuenta que 30 casos se asignaron forzosamente el criterio recién definido se utilizó para asignar 310 casos. Finalmente, se asignaron 60 casos más (de forma discrecional) para

¹ Simplemente por criterio de muestreo los tres grupos se construyeron de forma que sean excluyentes, primero se identificó el grupo de los queseros, a continuación, se identificó a todos los establecimientos colonos del INC que no se habían identificado previamente como queseros, y finalmente se identificó al resto de los remitentes (no identificados como colonos ni queseros).

aumentar la representación de los estratos 2 y 3; por todo esto la muestra original resultó en 400 casos distribuidos como muestra la columna 3 del cuadro 2.



Muestreo

El muestreo se realizó a partir del marco muestral definido en la primera sección, siguiendo un procedimiento aleatorio estratificado de acuerdo a los estratos antes definidos y listados en el cuadro 2.

En la práctica es habitual que algunas de las unidades seleccionadas en una muestra no puedan ser encuestadas por problemas para contactarlas por rechazos o por otras razones, por este motivo en cada estrato se seleccionaron más unidades de las que se deseaba encuestar, siempre manteniendo el criterio de selección aleatoria.

El cuadro 2 presenta la muestra final asignada al equipo de campo (columna 4). En total se asignó una muestra de 569 casos.

Finalmente, luego de la etapa de campo, se obtuvieron 359 respuestas válidas, tal como se muestra en la columna 5 del cuadro 2.

Cuadro 2. Unidades del marco, muestra, y respuestas por estrato

Estrato	N (población marco)	n ₀ (muestra de diseño)	n (muestra efectiva)	nr (muestra respondiente)
1-remitente	23	23	23	14
2-colono	27	7	12	8
2-quesero	7	6	7	2
2-remitente	218	41	56	41
3-colono	105	18	26	16
3-quesero	37	7	10	6
3-remitente	442	81	106	73
4-colono	225	16	24	18
4-quesero	194	15	27	13
4-remitente	625	60	79	55
5-colono	151	11	17	10
5-quesero	441	53	85	45
5-remitente	467	62	97	58
TOTAL	2962	400	569	359



Expansores muestrales

Siempre que se utilizan datos de encuestas es clave la utilización de expansores muestrales de modo que la muestra represente correctamente a la población de interés y sea posible hacer buenas estimaciones respecto a cantidades poblacionales. Para el cálculo de los expansores se siguió una estrategia en tres pasos según sugieren Valliant (2013) y Ferreira y Zoppolo (2017).

En un primer paso se calcula para cada estrato la probabilidad de un individuo perteneciente a ese estrato de haber sido muestreado. El inverso de dicha probabilidad constituye el expansor base o inicial, este es único por estrato y compartido por cada individuo de la muestra que pertenece a dicho estrato.

$$w_{0k} = N_h/n_h$$

Donde, w_{0k} es el ponderador inicial para el individuo k , N_h es el tamaño poblacional del estrato h , y n_h es la muestra efectiva del estrato h .

Cuadro 3. Expansores iniciales por estrato

Estrato	N	n	$w_0=N/n$
1-remitente	23	23	1,00
2-colono	27	12	2,25
2-quesero	7	7	1,00
2-remitente	218	56	3,89
3-colono	105	26	4,04
3-quesero	37	10	3,70
3-remitente	442	106	4,17
4-colono	225	24	9,38
4-quesero	194	27	7,19
4-remitente	625	79	7,91
5-colono	151	17	8,88
5-quesero	441	85	5,19
5-remitente	467	97	4,81

Luego de calculados los ponderadores iniciales se quitan de la muestra las unidades que en la etapa de campo se clasificaron como no elegibles por no seguir perteneciendo al rubro en el año de la encuesta (41 casos); de esta forma la muestra elegible asciende a 529 individuos y la muestra elegible respondiente a 357 individuos, estratificados según muestra el cuadro 4.

Cuadro 4. Muestra elegible y muestra elegible respondiente

Estrato	n muestra efectiva	n muestra efectiva
1-remitente	20	14
2-colono	12	8
2-quesero	6	2
2-remitente	56	41
3-colono	25	16
3-quesero	10	6
3-remitente	99	73
4-colono	21	17
4-quesero	25	13
4-remitente	72	55
5-colono	15	9
5-quesero	83	45
5-remitente	85	58
TOTAL	529	357

Como se adelantó y se desprende de la tabla 2 la no respuesta fue relativamente alta (la tasa global de no respuesta asciende a 36,9%). Por lo general la no respuesta depende de características (observables o no) de los individuos de la muestra, y difícilmente sea aleatoria, por lo tanto, los expansores de muestra que no tengan en cuenta esta característica darán lugar a estimaciones sesgadas.

Por ese motivo, en el segundo paso de la construcción de expansores se realizó un ajuste por no respuesta. Para ello se ajustó un modelo binario de probabilidad de respuesta en base a un vector de características de los individuos que están disponibles para todas las unidades elegibles, tanto para quienes responden como para quienes no responden la encuesta.

Posteriormente se crearon clases de acuerdo a los cuantiles de probabilidad estimados, y a cada individuo de su clase se le asignó la mediana de probabilidad estimada del cuantil². Finalmente, el factor de ajuste es el inverso de dicho valor.

El expansor de esta segunda etapa es el factor de ajuste recién detallado, multiplicado por el expansor inicial del primer paso.

$$W_{NRk} = g_{NRk} * W_{0k}$$

En tercer lugar, a los expansores de la segunda etapa se les realizó un ajuste que consistió en una calibración de tal forma que la muestra expandida estime de forma insesgada una cantidad conocida a priori para la población. Para esto, se volvió a estratificar la muestra en función de

² Otra opción es asignarle directamente a cada individuo la propensión estimada, aunque este procedimiento puede dar lugar a expansores muy variables o extremos cuando existen probabilidades estimadas muy pequeñas.

estratos discretos según cantidad de leche³ y según si se trata de un establecimiento remitente o no.

Para cada uno de los posts estratos definidos en esas dos dimensiones se calcularon factores de ajuste, en donde el factor de ajuste es el número de casos estimados en cada estrato en base a los expansores de la segunda etapa dividido el número de casos de la población en los estratos en función de información conocida para toda la población.

$$g_k = N_g / \hat{N}_g$$

Donde N_g es el conteo poblacional del post estrato, y \hat{N}_g es la estimación del post estrato g utilizando los ponderadores de los establecimientos ajustados por no respuesta.

Para este ejercicio de calibración se usó información poblacional a partir de una actualización del marco muestral incorporando información del año 2019 no disponible al momento de la estructuración del mismo⁴⁵.

Cuadro 5. Post estratificación

estrato	N _g		N̂ _g	
	Remite	No remite	Remite	No remite
menos de 108.500	327	294	200,8	289,9
entre 108.500 y 269.500	544	193	568,9	163,5
entre 269.500 y 489.500	460	74	426,6	32,7
entre 489.500 y 996.000	461	24	488	50,1
más de 996.000	428	13	493,7	2,2

Finalmente, el producto del expansor de la segunda etapa por el factor de ajuste de post estratificación es el expansor final de la muestra (w_k).

$$w_k = g_k * w_{NRk}$$

³ Los rangos para el post estrato de cantidad de leche son: menos de 108.500 litros, entre 108.500 y 269.500 litros, entre 269.500 y 489.500 litros, entre 489.500 y 996.000 litros, y más de 996.000 litros.

⁴ A los dos criterios filtro originales (elegibilidad geográfica y tamaño mínimo) se agregó el requisito de tener producción de leche positiva en 2018 y 2019.

⁵ Se quitaron del marco poblacional de comparación, dos importantes empresas no respondientes que por sus características particulares no son representables a partir de las unidades muestrales respondientes.



Variabilidad

Como forma de informar sobre los niveles de incertidumbre de la encuesta, se presenta la estimación de tres totales relevantes a partir de la encuesta final expandida y sus errores estándar. La tabla 6 muestra estimación y error estándar para el total de leche producida para la cantidad total de vacas masa y para la superficie total.

Cuadro 6. Totales estimados y errores estándar

	total	SE	coeficiente variación
Producción leche	1.788.160.680	59.515.318	3,3%
superficie	560.340	24.749	4,4%
vacas masa	358.003	10.207	2,9%



Trabajo de campo

La Encuesta fue realizada a través de entrevistas telefónicas realizadas a los números disponibles en DICOSE y en los casos en que no se contaba con los mismos, se buscaron otras fuentes para obtener los números. El trabajo de campo fue desarrollado por Opción Consultores y los encuestadores fueron técnicos vinculados al sector (veterinarios e ingenieros agrónomos). Los encuestadores recibieron capacitación en el llenado del formulario de la encuesta, impartida por el equipo técnico de INALE y de Opción Consultores (relacionado a cuestiones propias de formulario a utilizar).

Los datos relevados en el campo fueron posteriormente expandidos en función de la estratificación presentada anteriormente para obtener los datos poblacionales que son los resultados presentados en el cuerpo principal de esta publicación.

La Encuesta abarcó diversos temas con un grado de significativa profundidad en la que se relevaron tanto datos físicos de las explotaciones como aspectos tecnológicos. Cada encuesta se realizó por teléfono (en su mayoría) mediante el llenado de un formulario en la web que completó el encuestador en un celular, tableta o computadora. Las entrevistas tuvieron una duración promedio de dos horas. Fueron realizadas entre febrero 2020 y abril 2021.



Publicación de datos

En la página web del INALE se encuentran disponibles los datos obtenidos a partir de la Encuesta 2019 y los correspondientes a la Encuesta 2014. A los que puede acceder en: <https://www.inale.org/estadisticas/>, seleccionando el filtro Encuesta Lechera.